SYSTEM AND METHOD FOR GENERATING PSEUDO-RANDOM NUMBERS

Cross-Reference to Related Applications

The present applications claims priority to co-pending United States Provisional Patent Application No. 60/393,733 entitled "System and Method for Generating Pseudo-Random Numbers, filed on July 8, 2003, the entirety of which is incorporated by reference herein.

Background of the Invention

The present invention relates generally to the field of computer systems and, more particularly, to systems and methods for generating random or pseudo-random numbers within such systems, for the purpose of maintaining system security.

Many chips have the means to generate random numbers. These may be true random number generators, in which the randomness comes from some physical source such as shot noise or oscillator drift; or pseudorandom number generators, in which a key is used to generate a long sequence of bits that are hard to predict if the key is not known. Each has its advantages.

True random number generators produce outputs that are statistically independent of each other. Thus the compromise of some set of outputs, or of the state of the device at some given time, does not impair the security of random numbers generated in either the past or the future. On the other hand, true random number generators are tricky to design (many designs being dependent on the precise fabrication process); they are hard to test; they may be vulnerable to an opponent who can manipulate chip inputs (such as Vcc); and the rate at which random numbers are generated is usually fairly low.

Pseudorandom number generators can provide deterministic output at high rates; they can use thoroughly studied and well-understood building blocks; and can be made just as testable and resilient as the rest of the device. However their unpredictability depends on some cryptographic key remaining unknown to an opponent. While appropriate use of a one-way cryptographic function can prevent a key compromise being used to deduce previous inputs, there is no obvious way to recover security following a compromise. Unfortunately, known methods for generating such pseudo-random numbers such as encryption using the SHA-1 or DES algorithms do not afford the level of protection required to ensure that the cryptographic key remains secure.

Accordingly, there is a need in the art of computer systems for a system and method for generating pseudo-random numbers which overcome the security limitations of known systems.

Summary of the Invention

The present invention overcomes the problems noted above, and realizes additional advantages, by providing for methods and systems for generating pseudo-random numbers utilizing techniques of both the SHA-1 and DES encryption standards.

In accordance with one embodiment of the present invention a current seed value S_j is loaded from a non-volatile storage. Next, values E, representative of environmental randomness, and C, representative of configuration data are likewise loaded. A new seed value, S_{j+1}, is generated in accordance with the equation $S_{j+1} = f(S_j; A; C; E)$, wherein f represents a selected encryption algorithm , and B is a second constant, and wherein S_j is concatenated with A, which is concatenated with C which is concatenated with E. The

new seed is then written to the non-volatile storage. Next, a key, K, is generated in accordance with the equation $K = f(S_j; B; C; E)$, wherein B is a second constant. A pseudo-random number output, P_n, is then generated in accordance with the equation $P_n = f_{3DES}(K, P_{n-1})$, where f_{3DES} represents the operation of triple DES encryption hardware, and P_{n-1} is the previously generated pseudo-random number.

Brief Description Of The Drawings

The present invention can be understood more completely by reading the following Detailed Description of the Preferred Embodiments, in conjunction with the accompanying drawings.

FIG. 1 is a simplified flow diagram illustrating one embodiment of a method for generating pseudo-random numbers in accordance with the present invention.

Detailed Description of the Invention

Referring now to the Figures and, in particular, to FIG. 1, there is shown a simplified flow diagram illustrating one embodiment of a method and system for generating pseudo-random numbers in accordance with the present invention. In particular, the present invention utilizes key features of both of the above-identified methodologies. This combination of the two approaches results in a pseudo-random number generator that is re-keyed periodically using an external input of physical randomness.

The pseudo-random number output will be computed in step 108 using the 3DES (triple DES) encryption hardware, operated in output feedback mode. Writing $f_{3DES}(K, P)$ for the encryption of P using the key K, we have

$$P_n = f_{3DES}(K, P_{n-1})$$

Where, the initial value P_0 can be set to any fixed value such as 0. This will provide a source of pseudorandom numbers with a rate of about 15 Mbit/sec. The key K will be derived from a seed S kept externally in non-volatile memory. Initially, on power-up, the device loads the current value S_j of the seed, plus configuration data C and environmental randomness E in step 100. The device will compute the key K in step 102 and the next value S_{j+1} of the seed in step 104 as follows, using, in one embodiment, the FIPS 180 secure hash standard algorithm (SHA). The seed S_j will preferably be 160 bits in length if the current secure hash standard algorithm SHA-1 is used, and 256 bits if the proposed new standard SHA-256 algorithm is used:

$$S_{j+1} = f_{SHA}(S_j; A; C; E)$$

$$K = f_{SHA}(S_j; B; C; E)$$

Here A and B are two different fixed constants whose value is not otherwise critical (for example, A = 1 and B = 2). Furthermore, the phrase (X; Y) denotes X concatenated with Y. The configuration data C may be any length, and the environmental randomness E should have at least 80 bits of entropy.

Following this updating computation, the new seed value S_j is written back to non-volatile storage in step 106. The function of the environmental randomness E (which may be derived from whatever sources are available) is security recovery. If, at some time, the seed is compromised by an adversary who manages to read the off-chip non-volatile storage, the use of fresh randomness should deny him knowledge of subsequent seeds and the pseudorandom numbers derived from them.

In an additional embodiment, if it is desired to make it harder for a competitor to produce a compatible chip, then one or more of the components of the updating computation may be protected. A simple way to do this is to make A and B into secrets that are read from a ROM address that is not externally readable. A more thorough way to obscure the computation is to replace SHA with an algorithm that is proprietary. In the former case, A and B may be compromised by an opponent who mounts an invasive attack (such as microprobing the bus, or using a scanning capacitance microscope to read out the ROM). In the latter case, an opponent who performs this reverse engineering is further hindered by proprietary algorithm. If the threat of litigation is reckoned to be enough in itself, then, the constants A and B might contain as substrings the customer's copyright notice.

Further, if it is desired that a key compromised occurring during a session should not expose keys used earlier in that session, then the second equation of the key updating computation may be carried out more frequently than once per power cycle. In the limit, all the random numbers could be computed using SHA as successive values of the key K. In this case, as no use would be made of the 3DES hardware provided, there would be a noticeable performance penalty. Whether this mattered would depend on the application.

If it becomes a requirement at some future time to have a true random source on-chip, then this can be input to an on-the-fly update. Some care is needed though to ensure that enough random bits are input to each update that an opponent cannot work forwards by exhaustive search.

It should be understood that the 160-bit SHA-1 algorithm is in the process of being supplemented by the 256-bit SHA-256 and the 512-bit SHA-512 algorithms. Similarly, the existing standard DES modes of operation are in the process of being supplemented by the new dual counter mode. The practical consequences of these upgrades for technical security are few, but it may be decided to support them anyway in case they become a checkbox item for customers. For example, it might be objected that the output of SHA is only 160 bits, while 3DES uses a 168-bit key. In practice, the remaining eight key bits may be set to an arbitrary or zero value; but the objection is removed by the use of SHA-256 from whose output 168 distinct key bits may be drawn. Similarly, it may be objected that 3DES in output feedback mode will cycle after about 2^{32} pseudorandom values have been drawn; this is unlikely to be an issue in the envisaged applications, but the objection is removed by the use of the new dual-counter mode of operation, for which a further 64 bits of key is required to initialize the counters. In that case, one should use SHA-256 to provide the 232 bits required in total.

The pseudo-random number generation system of the present invention makes a number of assumptions about the physical protection of the equipment being protected. In particular, the present invention assumes that the protected device contains no on-chip non-volatile memory, thus requiring that any encryption key material must be stored off-chip. It follows that potential adversaries must not have unsupervised access to the equipment. In particular, the off-chip non-volatile memory is to be kept secure & inaccessible by unauthorised personnel. Further, it is assumed that the attacker does not

have unsupervised access to the electrical interface of the device or of associated chips, with which timing attacks might be possible.

While the foregoing description includes many details and specifities, it is to be understood that these have been included for the purposes of explanation only, and are not to be interpreted as limitations of the present invention. Many modifications to the embodiments described above can be made without departing from the spirit and scope of the invention, as is intended.